

МОДЕРНИЗАЦИЯ ЦЕНТРОВ ОБРАБОТКИ ОПЕРАТИВНОЙ ГИДРОМЕТЕОРОЛОГИЧЕСКОЙ ИНФОРМАЦИИ РОСГИДРОМЕТА

В.А. Анцыпович, С.В. Лубов

*ФГБУ «ГВЦ Росгидромета»
antsyp@hydromet.ru, loubov@hydromet.ru*

С целью выполнения национальных задач и международных обязательств Российской Федерации во Всемирной метеорологической организации (ВМО) в части численного прогнозирования в рамках Проекта модернизации и технического перевооружения учреждений и организаций Росгидромет провел переоснащение вычислительных подразделений.

Современный период развития вычислительной базы в Росгидромете характеризуется наличием трех уровней центров обработки оперативной информации:

- единый центр глобальной обработки информации в Москве, который выполняет и функции Мирового метеорологического центра в ВМО; в функции этого центра входят задачи сбора информации о состоянии окружающей среды в глобальном масштабе, их усвоение и расчет прогноза развития процессов по всему земному шару с различной (от 12 часов до 15 суток) заблаговременностью;

- специализированные центры, обеспечивающие расчет прогноза развития атмосферных процессов по своей зоне ответственности, мониторинг и прогноз развития распространения загрязняющих веществ, а также прогнозы изменения климата (центры Росгидромета в Новосибирске, Хабаровске, Санкт-Петербурге); выполняют функции региональных и мировых специализированных центров оперативной обработки гидрометеорологической информации; центр обработки

оперативной информации в Москве также выполняет функции регионального центра;

- локальные оперативные прогностические центры Росгидромета, обеспечивающие интерпретацию прогностической продукции, ее визуализацию, окончательную обработку и подготовку для конкретного потребителя; данные центры Росгидромета расположены по всей территории России.

Современный этап развития технологий прогнозирования атмосферных процессов для обеспечения повышения точности и заблаговременности прогнозов характеризуется переходом к расчетам по глобальным моделям атмосферы с горизонтальным разрешением 25–40 и менее километров и применением ансамблевых методов, связанных с одновременным счетом нескольких десятков вариантов моделей.

В условиях ограничения допустимого времени счета для выполнения задач усвоения оперативных данных и прогноза развития атмосферных процессов задачи глобального масштаба требуют производительности вычислительных комплексов в десятки терафлопс (1 TFlops = 10^{12} операций с плавающей запятой в секунду), а задачи регионального масштаба – производительности в сотни гигафлопс (1 GFlops = 10^9 операций с плавающей запятой в секунду).

В результате проведенных коллективом специалистов ГВЦ Росгидромета и Гидрометцентра России исследований, в 2005 году в Росгидромете был начат конкурсный процесс модернизации центров обработки оперативной информации в Москве, Новосибирске, Хабаровске и Санкт-Петербурге. Типовая структура модернизированного вычислительного комплекса для каждого из указанных центров представлена на рис. 1.

Вычислитель выполняет функции оперативного и научно-исследовательского счета, ресурсной поддержки научно-исследовательских разработок. Дифференциация задач выполняется по технологическим схемам каждого из центров установки. В комплект системы входит системная консоль организации доступа ко всем узлам системы для инженерного персонала.

Система управления данными состоит из:

- дисковой системы хранения данных, обеспечивающей хранение данных на накопителях ЖМД и предоставляющей ресурсы для подсистем ввода-вывода Вычислителя, серверов баз данных и оперативных систем;

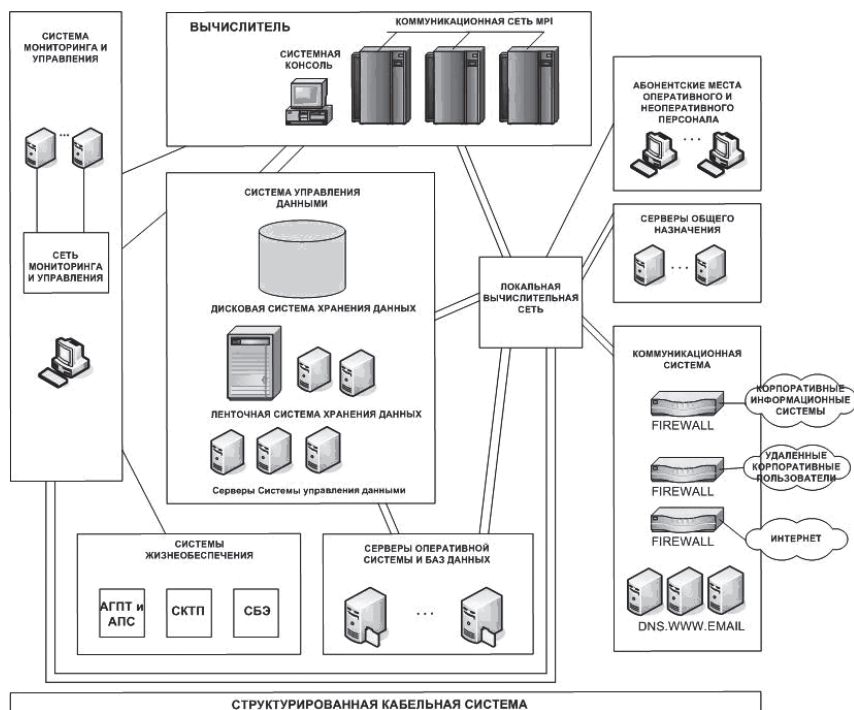


Рис. 1. Типовая структура модернизированного центра обработки оперативных данных Росгидромета.

- ленточной системы хранения данных, состоящей из автоматизированной ленточной библиотеки;
- серверов системы управления данными, включая серверы архивирования и резервного копирования, иерархического хранения данных.

Серверы оперативной системы и баз данных включают в себя вычислительные серверы, локальные серверы оперативного назначения, в число которых входят серверы обмена оперативной информацией, серверы информационных систем Росгидромета, файловые серверы, обеспечивающие файловый и транзакционный обмен данными для клиентов ЛВС.

Серверы общего назначения обеспечивают функционирование основных сетевых сервисов, таких, как DHCP, NTP, локальных DNS и SENDMAIL.

Абонентские рабочие места оперативного и научно-исследовательского назначения включают персональные компьютеры и станции визуализации оперативного и неоперативного персонала.

Локальная вычислительная сеть объединяет все компоненты информационной системы. Связь с существующей ЛВС осуществляется посредством двух интерфейсов Ethernet 10/100/1000 Base-T. Для обеспечения функционирования осуществляется мониторинг и управление сетевым оборудованием.

Система мониторинга и управления обеспечивает мониторинг и централизованное управление Вычислителем и Системой управления данными. Система мониторинга и управления предназначена для удаленного управления этими компонентами и включает в себя Сеть мониторинга и управления, которая объединяет специализированные аппаратные средства Вычислителя, Системы управления данными и Системы жизнеобеспечения.

Система информационной безопасности и удаленного доступа обеспечивает функционирование сервисов информационной безопасности и авторизации доступа к ресурсам, в том числе удаленного. В частности, к таким сервисам относятся системы антивирусной и анти-спамовой защиты, сервисы аутентификации и управления доступом, сервисы обеспечения безопасности ЛВС.

Структурированная кабельная система обеспечивает интеграцию всего активного оборудования вычислительного комплекса.

Конкурсный процесс был построен на условиях достижения максимальной производительности на основных вычислительных алгоритмах глобальных моделей атмосферы, применяемых в оперативной практике ММЦ в Москве, и минимальной стоимости контракта. В результате конкурса в 2007–2008 гг. были заключены соответствующие контракты и по результатам установлены вычислительные комплексы, основные характеристики которых указаны в таблице.

В специализированных региональных центрах в Новосибирске и Хабаровске, а также в Мировом центре радиационных данных в Санкт-Петербурге пиковая производительность Вычислителя компании

Таблица

Конфигурация информационно-вычислительных систем в центрах Росгидромета

Центр Росгидромета	Пиковая производительность TFlops (10 ¹²)	Тип Вычислителя	Количество CPU/CPU в узле	ОЗУ Тбайт (10 ¹²)	Производитель (фирма/страна)	Дисковая подсистема Тбайт(10 ¹²)
ММЦ Москва	11	SGI Altix4700	1664/128	6.6	SGI (США)	180TB=60TB SGI IS4000 + 120TB SGI IS10000
	16	SGI Altix ICE8200	1408/8	2/8	SGI (США)	
РСМЦ Новосибирск	0.7	G-Scale S4700	104/104	0.2	Крафтвэй (Россия)	12.3TB SGI InfiniteStorage4000
РСМЦ Хабаровск	0.7	G-Scale S4700	104/104	0.2	Крафтвэй (Россия)	12.3TB SGI InfiniteStorage4000
МЦРД (ГГО им. А.И. Воейкова, С-Петербург)	0.7	G-Scale S4700	108/108	0.2	Крафтвэй (Россия)	14.4TB Kraftway Storage 400S

G-Scale S4700 ЗАО «Крафтвэй Корпорэйшн ПЛС» на базе 54 процессоров Intel Itanium 9040 (для центра в Санкт-Петербурге) и 52 процессоров Intel Itanium 9140M (для центров в Новосибирске и Хабаровске) составляет 0,7 TFlops. Емкость дискового хранилища для центров в Новосибирске и Хабаровске – 12,3 ТВ на базе оборудования SGI InfiniteStorage 4000, а для центра в Санкт-Петербурге – 14,4 ТВ на базе оборудования Kraftway Storage 400S.

Основной вычислительный комплекс для выполнения глобальных усвоения данных наблюдений и прогноза состояния атмосферы установлен в Москве (рис. 2). Вычислительные мощности представлены двумя кластерами общей пиковой производительностью 27 TFlops. Рассмотрим на его примере построение вычислительных процессов подробнее.



Рис. 2. Фото Вычислителя SGI Altix4700, установленного в ГВЦ Росгидромета (Москва).

Кластер (Вычислитель #1) SGI Altix 4700 пиковой производительностью 11 TFlops построен на 832 двухядерных процессорах Intel Itanium 9140M с оперативной памятью 6,6 ТВ (1664 процессорных ядер, 4 GB на ядро). Вычислитель представляет собой систему с общей памятью, способен масштабироваться до 1024 ядер в рамках единого образа операционной системы. Архитектура SGI NUMAflex обеспечивает наилучшие (по проведенным тестированиям) значения полосы пропускания и латентности. Шина взаимодействия SGI NUMALink обеспечивает высокоскоростное взаимодействие между процессорами системы. На текущий момент Вычислитель логически делится на партии/узлы (13 узлов): 1 партия = 1 вычислительная

стойка. В каждой партии 32 виртуальных вычислительных узла (нод): 4 ядра + 16 ГБ оперативной памяти. Две виртуальных ноды зарезервированы для системных процессов (8 ядер + 32 ГБ памяти). Каждая аппаратная партия имеет системные жесткие диски SAS с полезным объемом 300 GB с зеркалированием и высокопроизводительный доступ к Дисковой системе хранения данных через параллельную файловую систему SGI CXFS.

Кластер (Вычислитель #2) SGI Altix ICE8200 пиковой производительностью 16 TFlops построен на 352 четырехядерных процессорах Intel Quad Core Xeon e5440 с оперативной памятью 2,8 TB (1408 процессорных ядер, 2 GB на ядро). Вычислитель представляет собой широко распространенный тип блейд-систем, имеет 64-разрядную архитектуру EM64T. Вычислитель состоит из трех вычислительных стоек, каждая из которых представляет собой: 4 IRU (Individual Rack Unit – блейд шасси), 15 вычислительных блейдов (лезвий), 8 ядер и 16 GB оперативной памяти в каждом блейде. В качестве шины взаимодействия используется технология Infiniband. Кластер использует бездисковый принцип работы – образ операционной системы разворачивается в оперативной памяти. Общая файловая система строится на базе технологии NFS over Infiniband. Таким образом, узлы кластера имеют только сетевой доступ к данным Дисковой системы хранения.

Дисковая подсистема общим объемом 180 TB имеет иерархическую структуру: 60TB SGI InfiniteStorage 4000 system и 120TB SGI InfiniteStorage 10000. Решение построено на сети SAN и файловой системе SGI CXFS, которая является разделяемой файловой системой, направляющей поток данных через SAN. CXFS обеспечивает соединение всех необходимых технологических серверов и узлов Вычислителя #1 с Дисковой системой хранения данных. Файловая система работает совместно с SGI DMF – системой иерархического управления хранением SGI Hierarchical Storage Management, которая обеспечивает миграцию данных между дисковыми подсистемами.

Для обеспечения единой точки входа для пользователей, организации среды трансляции организована система Front-end, состоящая из машин двух архитектур Intel Itanium и Intel Xeon соответственно. На вычислительных средствах установлено программное обеспечение,

включая компиляторы Intel для языков Fortran и C, научные библиотеки SGI SCSL, Intel MKL, библиотеки передачи сообщений MPI и SHMEM, отладчик Totalview.

Для обеспечения контролируемого запуска задач на ресурсах вычислительного комплекса организована система очереди задач на базе программного обеспечения Altair PBS Pro (рис. 3). С любой из систем Front-end допускается запуск задач на любом из кластеров в соответствии с имеющимися разрешениями. Рассматриваемая система имеет достаточно гибкую структуру и на настоящий момент организован запуск задач по приоритетам (задачи, имеющие статус оперативных, запускаются вне очереди и имеют возможность снимать с исполнения другие задачи меньшего приоритета), по размерности (небольшие задачи запускаются на ресурсах системы, которые обеспечивают их выполнение без останова) и т.д.

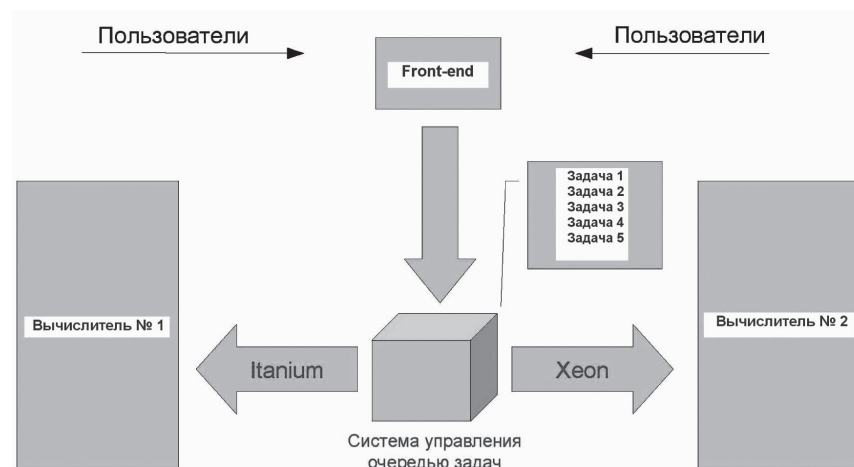


Рис. 3. Взаимодействие пользователей с системой.

Задачи модернизации и технического перевооружения, реализуемые Проектом Росгидромета, не завершаются поставкой конкретных программно-аппаратных комплексов. Большой объем работ по реализации прогностических алгоритмов и их развитию в среде с высокой степенью параллелизации был выполнен специалистами

Гидрометцентра России и ГВЦ Росгидромета в период 2003–2010 гг. К настоящему времени на вычислительном комплексе организовано выполнение прогностических расчетов на базе глобальных моделей атмосферы ПЛАВ-2008 и спектральной модели. Большим достижением стала организация оперативного счета по мезомасштабным моделям атмосферы COSMO-RU, WRF-ARW, требующим значительных ресурсов (800 процессоров для COSMO-RU, 700 процессоров для WRF). Идет интенсивный научно-исследовательский счет. Ресурсы комплекса существенным образом загружены. Особенно актуальным является задача дальнейшего развития вычислительной базы и увеличения производительности вычислительных комплексов с целью реализации на них новых оперативных прогностических технологий.

Опыт построения высокопроизводительных вычислительных комплексов и их системы жизнеобеспечения в вычислительных подразделениях Росгидромета показал, что центр сложности комплексного решения сейчас находится не в вычислительном комплексе, а в системе термовлажностной стабилизации, автоматический режим которой в экстремальных погодных условиях не совсем справляется с заданным функционалом.

При проектировании следующего этапа модернизации высокопроизводительных комплексов приходится учитывать, что на современном этапе развития мощностей каждые 2–3 года появляются решения, превышающие по производительности предыдущее в 20–25 раз и находящиеся примерно в том же самом энергопакете. Это означает, что через 2–3 года эксплуатация предыдущего комплекса становится невыгодной по параметру стоимости затрат на потребленное электричество. В свою очередь это означает, что невыгодно проектировать комплекс с «запасом» по мощности на длительный период времени, а следует ориентироваться на наращивание вычислительной мощности путем установки нового решения. Так как высокопроизводительные комплексы относятся к особо ценному имуществу с периодом эксплуатации не менее 10 лет, то в этой части также существует проблема в одновременном выполнении данного положения и законодательно предусмотренных мероприятий по энергосбережению.

При проектировании следующего этапа модернизации высокопроизводительных вычислительных комплексов Росгидромета, помимо задач портации существующих технологий, перед специалистами Росгидромета стоит задача анализа тренда развития вычислительных структур в ближайшие 3–5 лет и оценка эффективности реализации на них оперативных прогностических технологий. В апреле 2008 года руководство Росгидромета и компании Intel подписали протокол о создании Центра компетенции по высокопроизводительным вычислениям в области метеорологии. Центр компетенции призван объединить усилия ученых в области гидродинамического моделирования и специалистов информационных технологий для решения задач дальнейшего развития алгоритмов усвоения данных об окружающей среде и прогнозирования атмосферных процессов. В рамках Центра компетенции открываются возможности по исследованию будущих решений в области вычислительной техники.

Поступила в редакцию 15.07.2011 г.